Proceedings of the Fifth IASTED International Conference
VISUALIZATION, IMAGING, AND IMAGE PROCESSING
September 7-9, 2005, Benidorm, Spain

ELIN: A NEWSPAPER UNIVERSAL MULTIMEDIA ACCESS PLATFORM BASED ON MPEG STANDARDS

Jordi Casademont
Technical University of Catalonia (UPC)
C. Jordi Girona 1-3, Barcelona
Spain
jordi.casademont@entel.upc.es

Ferran Perdrix
Diari Segre
Carrer del riu 6, Lleida
Spain
fperdrix@diarisegre.com

Martin Einhoff
Fraunhofer Institute for Computer Graphics (IGD)
Fraunhoferstr. 5, Darmstad
Germany
Martin.Einhoff@igd.fraunhofer.de

Josep Paradells
Technical University of Catalonia (UPC)
C. Jordi Girona 1-3, Barcelona
Spain
teljpa@entel.upc.es

Georg Dummer
Forschungszentrum Informatik an der Universitaet Karlsruhe
Haid-und-Neu Str. 10-14, Karlsruhe
Germany
Dummer@fzi.de

Anne Boyer
Loria INRIA Lorraine Campus Scientifique
BP 239, Vandoeuvre-lès-Nancy, Cedex
France
Anne.Boyer@univ-nancy2.fr

ABSTRACT
Nowadays, technologies involved in Web information distribution services are evolving to adapt themselves to new user requirements. Usually, these new technologies are used separately. This paper presents ELIN (Electronic Newspaper Initiative) project, an European Commission funded project, that tries to integrate the newest standards and technologies involved in multimedia delivery. It has as objective the delivery of any type of media format to any kind of user terminal in an efficient way. In order to do that, it takes the approach of using MPEG standards: MPEG-4 for video delivery, MPEG-7 for data classification and MPEG-21 for data management and adaptation. So it integrates the totality of solutions provided for the MPEG group.
In order to test this technology, ELIN has chosen the on-line newspaper field. Therefore, it provides a toolkit to be used by mass media companies to set an advanced electronic newspaper. This toolkit enables to personalize the delivery of news using collaborative filtering, to build a XML file architecture to store and manage data, to deliver information in video (MPEG-4) or in HTML format for low capabilities devices, to adapt media for different capabilities terminals and includes a 3 Dimension environment to create an user community to communicate between them and share data objects.

KEY WORDS
Electronic newspaper, MPEG, adaptation, UMA, 3D, personalization.

1. Introduction

Web news publishing is evolving fast, as the majority of Internet services, and nowadays this service is trying to adapt information to a way that best fits user's interests in order to increase its usage. With that, newspapers are looking forward to take more profit from their news sites. In parallel, many of the newspaper companies are changing into news media houses. They own radio stations and video production companies that produce content unsupported by traditional newspapers, but that are able to be delivered by Internet newspapers. Also, the broadband access networks that users require to efficiently consume audio and video are spreading out. On the other side, with cellular phones or PDA (Personal Digital Assistant) users can reach news companies content from nearly any place of the world, although with several restrictions due to the low capabilities of their communication devices. This fact also give new challenges and opportunities for online newspapers that will be able to reach their users via SMS (Short Message Service) or MMS (Multimedia Message Service) in order to push important news.
We also have to point out that at the beginning, the news content of the news sites was the same as in the printed edition, while today there is a constant update of news during the day and users will be very interested on reaching this information as soon as possible [1].
Having all this issues in mind, we think that in order to have a good acceptation, Internet newspapers have to adapt themselves to the needs and profiles of its users. This adaptation can be seen from two points of view. The first one is adaptation on personal interests, or personalization. A newspaper may contain huge amounts of information. However, the average reader is only interested in fractions of the total amount and usually he has few time to read it. So, it is highly interesting to give the required news to each user according to his preferences, localization or simply the most striking news of that moment. The other point of view of adaptation is the data adaptation. We have already said that media companies can have the same news in different formats: video, audio or the traditional text plus images, also that

users can access this content from various type of devices (PC, PDA, smart phones) and using different bandwidth access networks. Therefore, the format of the data should be selected according to the system the user is accessing from, and if this data don't fits perfectly, it would require a further adaptation. These operation includes transcoding of formats, resizing of images or recoding for higher levels of compression.

Moreover, there are research working groups devoted to the definition of standards that can manage the whole framework of the consumption of media. It is the case of the MPEG (Moving Picture Experts Group). They have developed standards for coding audio and video: MPEG-1 (Video CD and MP3), MPEG-2 (Digital Television, set top boxes and DVD) and MPEG-4 (enables the production, distribution and interactive access of a flow of data formed by video, audio, images and graphics); also, they have elaborated MPEG-7, the standard for description and search of audio and visual content, and MPEG-21, the new standard that defines an open framework for multimedia delivery and consumption for use by all the players in the delivery and consumption chain.

In this context the European Commission, in its IST (Information Society Technologies) program, has funded a research project called ELIN: Electronic Newspaper Initiative. This project has built an electronic newspaper service that enhances consumer's experience by introducing interactive features, advanced personalization and ubiquity of usage. The other innovative aspect of ELIN is the global usage of MPEG standards: MPEG-4 for stream codification, MPEG-7 for data description and classification and MPEG-21 for data adaptation based on user's device capabilities. This project is a clear example of MPEG usage and shows its possibilities of interoperability. Although ELIN has been focused to be used as an electronic newspaper, it can be reconverted easily to any other media delivery service.

ELIN system has three main subjects: the application server, the content producer, and the consumer. The content producer site is provided by an authoring tool that easily enables journalists to build news in ELIN format using audio, video and text files. The consumer can access ELIN data using a MPEG-4 Media Player developed on the project, or simply using a HTML browser. With the first one, the user is able to access to the whole functionalities of the system while with the second there are several restrictions. The reason of enabling access with HTML browsers is due to the possibility that the user wants to get information from a low capabilities devices (PDA or smart phone) that don't support MPEG-4 data. In the server site run server applications that store media resources and their descriptors, control user access, personalize and adapt data, download media resources and send notifications of important news using e-mail, instant messaging, SMS or MMS. Finally, the system also provides a 3D environment communication tool. This 3D community environment enables the users to share data resources and to communicate between them with a secure videoconference system.

Having this considerations in mind, this paper is distributed as follows: section 2 presents the main topics of usage of MPEG standards, section 3 describes the XML structure of metadata information built in order to classify and store resource data into ELIN system, section 4 presents the used methodology for news personalization, section 5 presents how ELIN discovers the capabilities of user's devices and adapts data to fit them and finally section 6 describes the 3D community.

2. Usage of MPEG Standards

The Moving Picture Experts Group (MPEG) is a working group of ISO/IEC that is in charge of the development of international standards for compression, decompression, processing, and coded representation of moving pictures, audio and their combination. Although later, they found the need of other tools: the description of multimedia content and the interaction with content and players, so they developed two new standards, MPEG-7 and MPEG-21. ELIN system is one of the first frameworks to make use of the whole facilities provided by MPEG standards.

MPEG-4 is a standard, which describes binary streams for the efficient compression of visual and audio data. In addition, it specifies system issues for defining multimedia (combining natural and synthetic objects) interactive scenes (BIFS - Binary Format for Scenes), storing, multiplexing and transmitting the data and specifies how to interact with it. ELIN uses MPEG-4 to provide interactive news.

MPEG-7 (Multimedia Content Description Interface) provides a rich set of tools for completely describing multimedia content. In ELIN, we use MPEG-7 in order to store semantic metadata about news and user profiles. In addition, we decided to use IPTC (International Press Telecommunications Council) classification system [2]. IPTC is one of the most used classifications in the edition world. The MPEG-7 description allows to fill in all the necessary information about the creation, classification and composition for all news and advertisements.

Furthermore, the use of MPEG-7 description is an attractive feature for the Personalization Engine as it permits the access and filtering of different data sources using XML tools. Also it will provide a strong compatibility with other filtering platforms that use this standardized classification.

MPEG-21 is an open standards-base framework for multimedia delivery and consumption. It aims to enable the use of multimedia resources across a wide range of networks and devices [3].

MPEG-21 is based on two essential concepts: the definition of a fundamental unit of distribution and transaction, the Digital Item (DI), and the concept of users interacting with DIs. In order to do that, MPEG-21 is, at the time of writing, divided in 16 parts [4]. In particular, part 7, also called DIA (Digital Item Adaptation) takes into account all issues related with terminal adaptation, session maintenance an user localization.

MPEG-21 DIA [5] defines a set of descriptors and tools to enable transparent use of multimedia resources across different networks and devices using media adaptation. In the context of MPEG-21, a Digital Item is defined as a structured digital object (a XML file) with a standard representation that can describe a user or a multimedia resource. Nevertheless, MPEG-21 DIA takes care of the definition and structure of such a descriptors but not on how the values are acquired or data is adapted for different terminals. The basic idea is to have different types of DI describing groups of parameters:

– User preferences, terminal and network capabilities, and natural environment characteristics where the media is displayed.
– Parameters about data resources: format information and MPEG-7 meta-information (inserted into the DI) to classify the media.
– Parameters for session mobility.

ELIN uses MPEG-21 for news classification and terminal adaptation. For the first function, each news, add and media resource has been assigned a XML file named Content_DI, that using MPEG-7 meta-information describes and classifies this item (more details in next section). The Content DI will be subdivided in two groups: the Media DI and the News DI, this classification will be discussed in next section. All Content DI do not change during runtime.

For adaptation and personalization, each user has been assigned another XML file named Context DI. This one includes information about user preferences and capabilities of the terminal and network from where the user is accessing the system. Also if user's terminal has a GPS (Global Positioning System) receiver, the system will know his position and store it in the user's Context DI. This Context DI, that in fact has session information, changes during runtime.

In order to update Context DI, user's terminal send information to the server using the so-called DIA fragments. These are XML files, conformant with MPEG-21 DIA standard, that only contain a part of all possible parameters.

Therefore, Content DI and Context DI reside in the server, they are stored using a XML database, and DIA fragments are transmitted over the network to communicate the server and the client.

3. Content storage

In an editorial office there are a lot of applications creating media in several formats. To solve the problem of format integration, we need a structure that allows merging in the same schema different formats for each media.

Also, we want to consider only a single object for each unit of information, in this case for each news. So, when searching something upon this structure, all related content comes together.

Another interesting issue is that news can be linked to other news. This is an important issue taking into account the users' behaviour in terms of search. Usually when users search something they like to be given a group of related news. This link between news allows creating information threads.

Into ELIN system we have created a metadata file system structured in three hierarchical levels (figure 1). The lower level consists on the *physical files*, in whatever format they are, that contain the media data to be delivered to the user. They are stored in the file system in their initial format. The middle level, is formed by XML files that contain descriptors (what, when, where, how, who is involved, author, etc.) for each media object, one media object can be formed of several files of the same information resource but in different formats. We call these XML files *media digital items* (Media DI) that in fact are a type of Content DI. Finally, the top level consists of other XML files that describe the news by themselves and identifies its media components. We call these *news digital items* (News DI). There is one News DI for each news of the system. As an example, think on one news defined by a News DI named *news1.xml*, this file contains the descriptors of this news and identifies its components, lets say *text1.xml*, *video1.xml*, *image1.xml* and *image2.xml*, all these files are Media DI and contain descriptors of their media and the path of the file system where the physical file can be found. All this DI, according to MPEG-21 DIA structure are grouped into the Content DI group. ELIN system stores all this DI into a XML database.

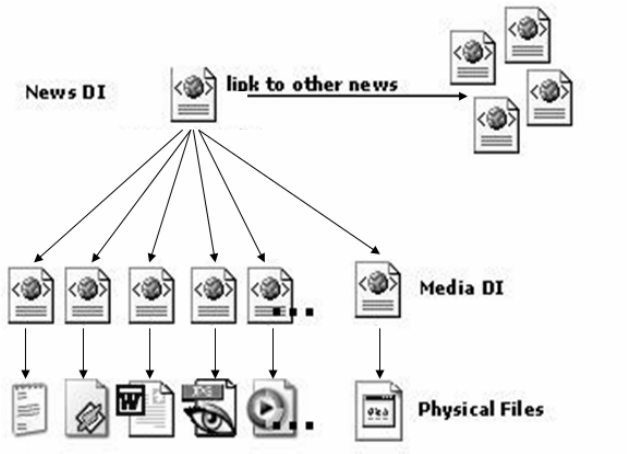

Figure 1. Content Digital Item structure.

The proposed schema allows reusing each media without limit. A Media DI can be part of different News DI, in this way a data resource can be accessed from multiple news. It is also possible to classify or index a single Media DI or a News DI. This is possible because all the digital items contain their own metadata. This metadata describes all the relevant characteristics for each case.

Using XML definitions, we improve the model interoperability and extensibility. This makes easier to add a new format and more metadata fields. Also, it is possible to import information from other archives or to discover data from an unknown source.

4. News Personalization

Considering information retrieval, powerful search and navigation mechanisms are needed for information consumers to find their way through. Current technology typically provides keyword-based search (often by fields: body, headline, section, lead, byline), browsing facilities inside newspaper issues, and, in online newspapers, navigation through static hand-made hyperlinks between news materials (e.g. links to earlier background stories).

ELIN system implements a filtering solution into a module named Smart Agent to search and personalize news. A smart agent has been defined by Nwana [6] as an autonomous collaborative learning interface agent. It means that it is goal-directed, proactive, and able to learn and improve with experience. It can work in concert with other agents to achieve a common goal. It is user adaptive. By adaptive, it means that such systems attempt to modify their behavior to maximize the productivity of the current user's interaction with the system. Obviously, for any such adaptation to be viable, it must be based upon information gathered concerning the current user behavior. Taking this in main, ELIN filtering system is based on two kinds of algorithms:

• A filtering algorithm that will be called "SIFA" (for Simple IPTC Filtering Algorithm) based on different levels of filtering.

• A Collaborative Filtering Algorithm which will be called CFA.

The SIFA algorithm works in two steps, each combining different filters. First, SIFA checks the IPTC preferences in the user profile and matches it with the IPTC description of documents. Then it keeps only the new documents (a new document is a document that has been published within the last x days, where x is a parameter of the filter). The second step is working on geographic position. If the document is "localized", it means that its description includes localization information, the Smart Agent takes into account the position of the user stored in his Context_DI to select or reject the resource. The Smart Agent sets a gap between localization in the Content_DI description of documents an the localization in the Context_DI of the user if available. This gap is a parameter in the Smart Agent setup. Finally, the smart agent checks the localization in the user profile and matches it with dissemination information in the Content_DI description of documents.

We have to notice that the quality of the SIFA filtering is highly related to the quality of the indexing provided by the publisher and the correctness of the static profile defined by the user. If localization and IPTC codes are not set correctly, the selection of pertinent resources will not be relevant. In fact, the SIFA filtering relies on the basic idea that a comparison between the description of a resource and the profile of a user in terms of topics of interest is sufficient to filter relevant documents.

The collaborative algorithm does not use the indexing of a resource nor the static profile of a user to select or reject a resource: it relies on an analysis of the usage. In practical terms, the collaborative filtering algorithm amounts to identifying the present user with a group of persons having the same tastes. The principle lies in the fact that users having appreciated the same documents have same interests. Then, it becomes possible to take advantage of the experience of a similar population to identify data likely to live up to reader's expectations.

The Smart Agent architecture allows the use of any collaborative filtering technique, if packaged using the agent specifications we provide. During the test phase of the ELIN project, we propose to use the Maes filtering algorithm implemented using the Pearson distance. This collaborative filtering algorithm is a memory-based algorithm: it means that it maintains a database containing votes of all users. A similarity score is determined between present user and each of the others members. Then, each prediction leads to a calculus on all of this source of data. The influence of a person is all the stronger in it since his/her degree of similarity with the present user is great. This memory-based technique offers the advantage to be very reactive, by integrating immediately modifications of users profiles into the system and is totally independent from the way the users has registered his/her profile.

Indeed, a collaborative filtering system evolves and learns in the course of time, according to actions done by users. So, it provides people with pertinent data, only when it has collected enough information on them. These kind of systems begin to be operational after a certain working period which depends on the number of users and resources.

5. Media Adaptation

As it has been said before, ELIN system is intended to be used in any terminal (static or mobile) and using any kind of access network (fixed or wireless). The simplest case is when the user accesses from a broadband fixed network using a standard PC. But nowadays the behavior of the users is evolving and they wish to access information from every place they are. In order to enable this, new broadband wireless access networks are appearing. Third generation (3G) mobile networks systems and IEEE 802.11 based networks offer high data rates to a wide range of mobile devices (laptops, PDAs, mobile phones). Furthermore, the apparition in next years of HSDPA (High Speed Downlink Packet Access) and mobility support in IEEE 802.16 (Wireless MAN) standards into commercial systems will expand the mobile market allowing users to access broadband wireless services from anywhere, anytime and anyhow. On the other side, mobile devices usually have a variety of singular hardware (screen size, keyboard,) and software capabilities (audio/video codecs, Java, operating system,…).

Due to the great variety of devices and means of connection the system has to implement a way of data adaptation. This process includes a functionality of data adaptation itself an another of discovering terminal capabilities.

The additional capacity of knowing the capabilities of the terminal have been identified over the last few years by different entities. The first to identify the problem was the W3C (World Wide Web Consortium) with the corresponding definition of the CC/PP (Composite Capabilities / Preferences Profiles). This initiative has remained practically unused due to the lack of a vocabulary. Other initiatives such as the one coming from the OMA (Open Mobile Alliance) called UAProf obtained better acceptance because it deals with a clear problem with the mobile devices and describes a vocabulary that facilitates the adoption by the browser manufacturers and Web content providers. In addition, the IETF (Internet Engineering Task Force) has identified this problem and created the CONtent NEGotiation (CONNEG) working group that finished its work in October 2000. The MPEG consortium has also developed a proposal oriented mainly to moving image. This initiative is enclosed in the MPEG-21 set of recommendations, in particular identified as DIA (Digital Item Adaptation). The MPEG-21 DIA results are useful for defining objects being exchanged, and extends the medias supported by W3C, but does not define any transport mechanism for information exchange between the provider and the consumer and does not support device identification. The W3C recently released a new document, entitled "CC/PP Structure and Vocabularies 1.0 Specification for Mobile Devices". This recommendation is an attempt to accommodate existing UAProf profiles. CC/PP is compatible with IETF media feature sets (CONNEG, RFC2533) in the sense that all media feature tags and values can be expressed in CC/PP. However, not all CC/PP profiles can be expressed as media feature tags and values.

In ELIN system the MPEG-21 DIA approach is used. The procedure of transferring terminal capabilities to the server is mixed with the registration process and works as follows. When a user wants to access the system he uses a registration client that communicates with the system, and registers the user after having validated his login and password. This interchange of information is done using SIP (Session Initiation Protocol - RFC 2543). The SIP protocol is an application-layer control (signaling) protocol for creating, modifying and terminating sessions with one or more participants. SIP invitations used to create sessions, carry session descriptions, which allow participants to agree on a set of compatible media types.

When the user has been validated, the next step is to communicate terminal capabilities to the server. For that, the terminal will send a XML file, following the structure of a MPEG-21 DIA fragment, to the server using a TCP/IP connection. This file will include all terminal capabilities and user's geographic position in case it is known by the terminal. The problem at this point is how this file is built, because it is not efficient that the user writes in its terminal parameters.

The solution taken in ELIN is that the registration client, if it is unable to find the capabilities file, will launch a Internet browser to a specific URL. Doing so the system can acquire all parameters sent in the HTTP (Hyper Text Transfer Protocol) header of the request (char set, language, encoders, browser type and version, etc.) and then it downloads a JavaScript applet that will discover the rest of required terminal capabilities (screen width and height, characters width and height, pixel depth, brand and type of terminal in case of smartphones, etc.) and network bandwidth making a measure of a file downloading speed. With all this information the server will build the terminal capabilities XML file and using a TCP/IP connection will send it to the terminal registration client that in turn will store the file in a specific directory for the next time the user registers to the system. With this methodology it is intended that each user terminal has stored its description file that will be sent always in the registration phase. The procedure of creating the file will only be done once for each terminal.

Another kind of capabilities interchanging procedure to have in mind for a future ELIN prototype is the possibility to store the MPEG-21 descriptions in the server and just to send a pointer from the terminal.

The Adaptation Server is divided in two parts according to the adapted type of data: video/audio streams and images. The data will be adapted according to the terminal and network capabilities learned by the system during the registration phase and stored into the Context DIs.

Video and audio streams adaptation: The system will have variations of the same information pre-coded at different rates, with different compression factors. Depending on the terminal and network capabilities and user preferences, the server will download one or the other. This adaptation follows a worst-case scenario and will use the minimum match (for instance, if 80/120 Kbps versions are available and the terminal supports 100 Kbps the 80 Kbps version is used).

Image adaptation: Images will be downloaded in the way that best fit terminal and network capabilities and user preferences (download speed versus image quality). Image adaptation will be done in real-time, taking into account: file format (jpeg, gif, png, bmp and wbmp), image resolution, number of color depth bits and compression factor.

6. 3D community

The ELIN 3D tool is developed as a shared virtual community environment for the PC platform where the different types of multimedia data are represented as textures on virtual walls/panels or as 3D objects in a virtual newspaper world. Beside this, 3D user interface serves as an integration platform for underlying components and services like Internet browser, MPEG-4 player, chat and audio/video-communication.

The virtual newspaper world consists of a basic scenery, whose appearance can be tailored to the respective target audience by newspaper publishers. This basic scenery can be arranged into subject areas just as actual newspapers are, for instance by placing different buildings within the virtual world. News feeds can be projected on multimedia canvases inside and outside these buildings.

Figure 2. A newspaper reader reads the forecast.

The virtual newspaper world is modelled in a way that allows for an intuitive navigation and enables the user to quickly and easily access those news feeds that are of particular interest to him. The graphical user interface allows for visualization and enrichment of news by means of three dimensional objects.

Within the virtual environment, avatars, i.e. computer-animated characters, represent readers of an electronic newspaper (figure 2). These avatars enable the user to navigate through and interact with the surrounding virtual world. When the reader moves its avatar, he can see and communicate with other users' avatars, too.

A directly integrated MPEG-4 based high quality audio/video communication service enables the readers to easily and intuitively talk to each other, whereby individual communication streams are visualized directly at the communicating avatars (figure 3) [7].

Figure 3. Two readers talk about the car with the integrated audio/video conferencing system.

A group of readers can jointly experience the multimedia information and three dimensional objects. ELIN pushes the boundaries of simple visualization: each user can interact with 3D objects and presentation screens, in order to show and discuss specific news with other avatars while he is visible to all other users. The 3D environment always shows each connected user the same state at the same time. With this functionality, the 3D environment offers a shared space between all users currently present within a 3D scene. All components hosted by the 3D environment like Internet browser and MPEG-4 player with the underlying news content are normally not shared between geographically distributed users.

7. Conclusions

This paper has presented a toolkit that integrates the newest developed standards and technologies in multimedia delivery in order to build an on-line newspaper. One of the main objectives has been to make use of MPEG standards: MPEG-4 for video delivery, MPEG-7 for data classification and MPEG-21 for data management and adaptation.

This toolkit enables to distribute data in video format (MPEG-4) and HTML (this case will be used only for low capabilities terminals). Data can be displayed using media players or inside a 3D environment that can be used also to communicate with other users. The system has a smart agent, used to personalize news, based on user profiles and collaborative filtering, in that way we select the most important news for each user. Information is adapted for user's terminal: standard PCs, PDA or also smartphones are considered. In order to adapt information a device capabilities discovering procedure plus a registration phase using SIP are implemented. Finally, the toolkit has an alert system that can send important notifications to clients using an instant messaging system, e-mail, MMS or SMS.

8. Acknowledgements

We want to acknowledge the other partners that have participate in the project: Grupo Alamo, Linkoepings Universitet and Aktiegolaget Oestgoeta Correspondenten.

This work was funded by the Commission of the European Communities through its Information Society Technologies Programme and the CICYT project TIC2003-01748.

9. References

[1] L. B. Eriksen, C. Ihlström, Evolution of the Web News Genre – The Slow Move Beyond the Print Metaphor, *Hawaii International Conference on Systems Science 2000*, CD-ROM.

[2] IPTC Council, http://www.iptc.org.

[3] I. Brunet, R. Van de Walle, K. Hill, J. Bormans, F. Pereira, MPEG-21: Goals and Achievements, *IEEE Multimedia*, Vol. 9, No. 2, October-December 2003, 60-70.

[4] http://www.chiariglione.org/mpeg/working_documents.htm

[5] MPEG-7: ISO/IEC 15938-7: Conformance, International Standard

[6] H. Nwana, D. Nduma, L. Lee, J. Collis, ZEUS: a toolkit for building distributed multi-agent systems, *Artificial Intelligence Journal*, Vol. 13, No. 1, 1999, 129-186.

[7] M. Einhoff, Desktop videoconferencing in distributed environments, *3rd International Workshop on COmmunication TOols for a Global Accelerator Network (CoToGAN 03)*, Trieste, Italy, October 28-31 2003.